C33

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : G06F 9/46

A1

(11) International Publication Number: WO 94/03855

(43) International Publication Date: 17 February 1994 (17.02.94)

(21) International Application Number: PCT/FI93/00304

(22) International Filing Date: 28 July 1993 (28.07.93)

(30) Priority data:
923430 29 July 1992 (29.07.92) FI

(71) Applicant *(for all designated States except US)*: NOKIA TELECOMMUNICATIONS OY [FI/FI]; Mäkkylän puistotie 1, FIN-02600 Espoo (FI).

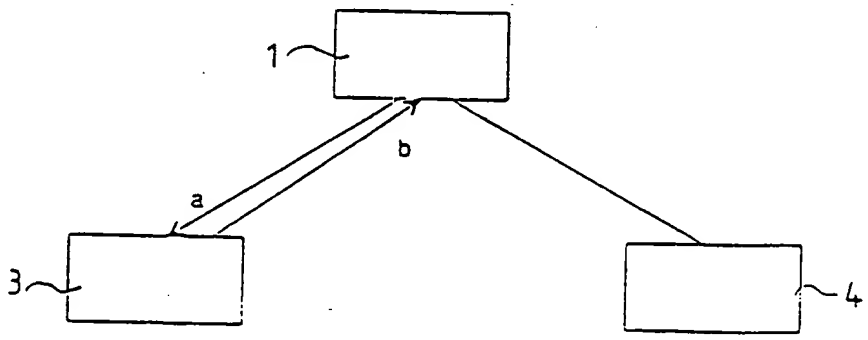
(72) Inventor; and
(75) Inventor/Applicant *(for US only)* : SYVÄNIEMI, Martti [FI/FI]; Nissnikuntie 1 B 8, FIN-02430 Masala (FI).

(74) Agent: OY KOLSTER AB; Iso Roobertinkatu 23, FIN-00121 Helsinki (FI).

(81) Designated States: AT, AU, BB, BG, BR, BY, CA, CH, CZ, DE, DK, ES, FI, GB, HU, JP, KP, KR, KZ, LK, LU, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SK, UA, US, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: METHOD FOR MANAGING RESOURCES ALLOCATED IN A COMPUTER



(57) Abstract

The invention relates to a method for managing resources allocated in a computer, wherein the allocated resources are maintained or released by time monitoring. In the invention, a task controller (1) is allocated for the management of resources required for a task to be executed; the controller (1) allocates resources (3, 4) required to execute the task; the controller starts time monitoring in each resource; if required by the task, the controller renews the resource allocations before the expiry of the respective time monitoring periods; and the controller (1) releases all resources allocated for the task after the completion of the task.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NE	Niger
BE	Belgium	GN	Guinea	NL	Netherlands
BF	Burkina Faso	GR	Greece	NO	Norway
BG	Bulgaria	HU	Hungary	NZ	New Zealand
BJ	Benin	IE	Ireland	PL	Poland
BR	Brazil	IT	Italy	PT	Portugal
BY	Belarus	JP	Japan	RO	Romania
CA	Canada	KP	Democratic People's Republic of Korea	RU	Russian Federation
CF	Central African Republic			SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovak Republic
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	UA	Ukraine
DE	Germany	MG	Madagascar	US	United States of America
DK	Denmark	ML	Mali	UZ	Uzbekistan
ES	Spain	MN	Mongolia	VN	Viet Nam
FI	Finland				

Method for managing resources allocated in a computer

This invention relates to a method for managing resources allocated in a computer, wherein the allocated resources are maintained or released by time monitoring.

The execution of a specific task in a computer system often requires co-operation between several resources based on recurrent use of the hardware, software and/or memory resources. The use of resources thus refers to the allocation of the parts of the computer system for the execution of a specific task, no matter whether these parts belong to the hardware, computation or memory capacity of the computer.

The resource allocation system has been problematic in that the known computers do not comprise any particularly advanced systems for monitoring allocated resources. Especially after the completion of the task assigned to the computer, the release of all or some of the resources may fail for some reason, e.g. due to malfunction, a programming error, etc. In practice, this problem easily results in that the system runs short of certain kind of resources in the long run, or that there are needlessly allocated resources in the system for long periods of time. This in turn reduces the capacity and may finally cause the operation to cease if all resources of a certain type have been allocated for one or more tasks which have been completed a long time ago.

One solution to the problem is offered by fixed-time monitoring which can be used in cases where it is possible to determine a maximum duration for the task to be executed by the resource. The resource is allocated for the maximum period of time, after which the resource releases itself if it has not yet been

released. Periodic monitoring has been used for resources allocated for an indefinite period of time or for resources allocated for a very long period of time. It is thereby checked intermittently that all resources in the system have changed tasks. If this has not happened, it is checked whether the main task is still uncompleted throughout the system, on the basis of which the resources are then released or maintained unchanged. One application of this procedure is the virtual memory management, in which the central memory is divided into pages (resources), and the use of different memory pages is monitored during the execution of a program. Unused pages are replaced with new pages from a mass memory whenever required. A disadvantage of periodic monitoring is that it loads the system and is complicated, as the monitoring procedure that checks all allocated resources has to be started intermittently, and if the system has to operate in parallel with the checking routine, a very complicated management algorithm will result as new resources are allocated all the time while old ones are released. In addition, a priority conflict occurs in the fixed-time monitoring of long tasks between the monitoring frequency and the execution if the average required allocation time is only a fraction of the time required for the entire task so that there occur needlessly allocated resources within the time monitoring. For instance, the memory pages to be replaced in the virtual memory management are not released until required, i.e. the central memory may, in principle, be filled up, which reduces the capacity of the system e.g. during the execution of background routines which do not use the virtual memory.

The object of the present invention is to solve the resource management problems by eliminating

resource allocations in cases where the resource remains needlessly allocated for an indefinite period of time. To achieve this, the method according to the invention is characterized by

allocating a task controller for the management of resources required for a task to be executed;

allocating by the task controller resources required to execute the task;

starting time monitoring in each resource by the controller;

renewing the resource allocations by the controller before the expiry of the respective time monitoring periods, if required by the task; and

releasing by the controller all resources allocated for the task after the completion of the task.

The task controller according to the invention thus operates so that when it allocates a resource for the execution of a specific task, it also requests the resource to start the time monitoring itself for a period of time which is either equal to the maximum allocation time when this suits the nature of the task, or with a resource to be allocated for an indefinite period of time, the allocation time depends on the distribution of use so that e.g. 90% of all cases (on a statistical basis) are processed within the allocation time. The method according to the invention is suitable for the management of both hardware resources (HW) and software resources (SW, e.g. memory buffers). The method is very reliable in operation as its operability is based on time monitoring, and all error situations can be handled simply by a restart. The method according to the invention is also operative in an environment in which communication between different resources is unreliable. The

method supports well the modern modular and object-oriented programming methods.

Other preferred embodiments of the method according to the invention are characterized by what is disclosed in the accompanying claims.

In the following the invention will be described by means of examples with reference to the attached drawings, in which

Figures 1a to 1c illustrate the resource management sequence of the method according to the invention in a normal case;

Figure 2 illustrates the renewal of a resource allocation; and

Figures 3a to 3c illustrate various resource allocation release situations.

In Figure 1a, a computer system, such as a central unit 2, i.e. the external environment from the viewpoint of a task to be executed, allocates a controller 1 for the task. The controller may be a hardware or memory area resource capable of controlling other resources in a co-ordinated manner. The controller may be scheduled to release itself spontaneously after the expiry of a fixed time if the external environment 2 does not separately renew it, or it may be releasable by the external environment 2 only.

In Figure 1b, the controller 1 allocates required resources from the computer system; in this specific case, it allocates a read circuit 3 and a memory buffer 4 of a keyboard (not shown) as hardware resources. The task assigned to the controller is thus to read characters from the keyboard to the memory buffer. Moreover, the controller starts a separate time monitoring in the keyboard read circuit and memory buffer 4. The time monitoring period may vary

from one resource to another, being e.g. 60 seconds in the read circuit 3 and 120 seconds in the memory buffer 4. As mentioned above, the resources can be allocated on the basis of the maximum time reserved for the task or on the basis of the probable task duration.

After this the time monitoring of the resources does not strain the controller or any other part of the computer system but each resource monitors itself in this respect. Time monitoring can be realized e.g. by a simple time counter program which is attached to the program controlling the operation of each resource or is activated in the program or hardware resource.

In Figure 1c, the system executes a given task by using the allocated resources, and after the task is completed, all resources are released immediately. Communication with the resources takes place via the controller 1. If the controller has been allocated for a fixed period of time, its possible renewal has to take place here, prior to the release of the controller's own time monitoring.

Figure 2 shows a case in which the renewal of a resource is required to complete the task. If the resources allocated in Figure 1b have been allocated for 50 seconds, and the task has not yet been completed, the controller 1 renews the allocation of the read circuit 3 by a renewal request a for another 60 seconds. The buffer 3 thereby acknowledges (arrow b) the further allocation to the controller, or does not accept/does not respond to the renewal request so that the controller can also release itself.

The hand-shaking according to the invention ensures that the resources remain needlessly allocated for no more than the duration of the time monitoring of the particular resource. The resource is released

on the basis of the time monitoring only when it is not released by the controller after the completion of the task, i.e. when the resource might otherwise remain allocated for an indefinite period of time; the renewal of allocation is required only seldom, depending on the given use distribution parameters, e.g. in 10% of all cases. Accordingly, the resource management binds very little capacity in the method according to the invention.

The minimizing of the duration of the time monitoring and the minimizing of the number of renewal requests are mutually contradictory objectives depending on the reliability of the system and the available capacity, whereby the required compromise is determined by parametrization, i.e. fixed parameters specific to the central unit, or computed parameters optimized for each task, for instance.

Correspondingly, the allocation of the memory buffer 4 in Figure 2 is renewable, if required, after 110 seconds from the start of the task. After the renewal of the allocation one returns to the situation shown in Figure 1c to wait for the completion of the task.

Figure 3a shows a situation in which a system part (in this specific case the read circuit 3) is damaged and/or incapable of communication. The following renewal request a of the allocation of the task controller 1 thereby fails (the acknowledgement b will not be received), and so the controller 1 releases itself, possibly after having sent an indication c of the failure of the task to its starter (the central unit).

In the situation shown in Figure 3b the fixed-time monitoring at the memory buffer 4 is released, and the memory area is freed. This situation typically

results from the hardware failure described in connection with Figure 3a or from the malfunction of the controller. Accordingly, all resources are released when the respective time monitorings are released even though the communication between the resources would be unreliable.

Figure 3c shows a case in which the task controller fails or its allocation for a predetermined time is not renewed so that it is released. Thereafter all other allocated resources are released in accordance with their own time monitoring procedures if there are no renewal requests.

It is obvious to one skilled in the art that the different embodiments of the invention are not limited to the illustrating embodiments described above but they may vary within the scope of the attached claims.

Claims:

1. Method for managing resources allocated in a computer, wherein the allocated resources are maintained or released by time monitoring, c h a r a c t e r i z e d by

allocating a task controller (1) for the management of resources required for a task to be executed;

allocating by the task controller (1) resources (3, 4) required to execute the task;

starting time monitoring in each resource by the controller;

renewing the resource allocations by the controller before the expiry of the respective time monitoring periods, if required by the task; and

releasing by the controller (1) all resources allocated for the task after the completion of the task.

2. Method according to claim 1, c h a r a c t e r i z e d in that the task controller (1) is formed as a resource provided with time monitoring.

3. Method according to claim 1 or 2, c h a r a c t e r i z e d in that when the time monitoring period of each resource (1, 3, 4) expires without a renewal, the resource is released spontaneously.

4. Method according to claim 1, 2 or 3, c h a r a c t e r i z e d in that the time monitoring period of each resource (1, 3, 4) is determined on the basis of the maximal execution time of the task.

5. Method according to claim 1, 2 or 3, c h a r a c t e r i z e d in that the time monitoring period of each resource (1, 3, 4) is determined on the basis of the probable execution time of the task.

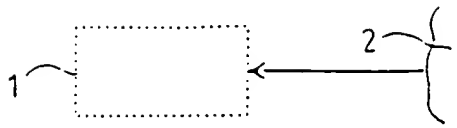


FIG.1a

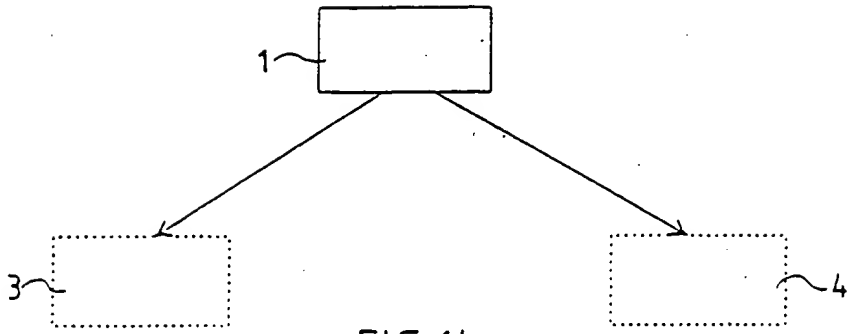


FIG.1b

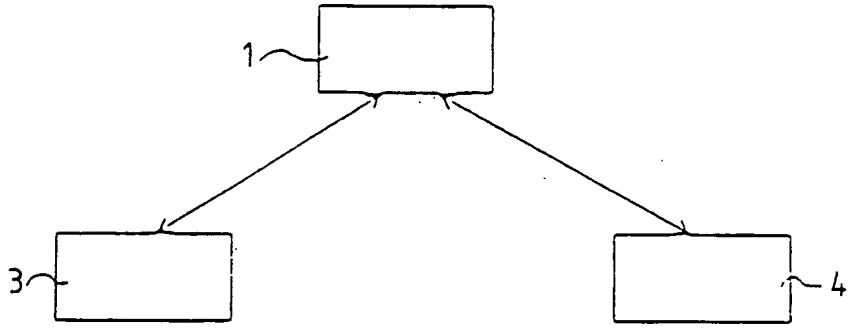


FIG.1c

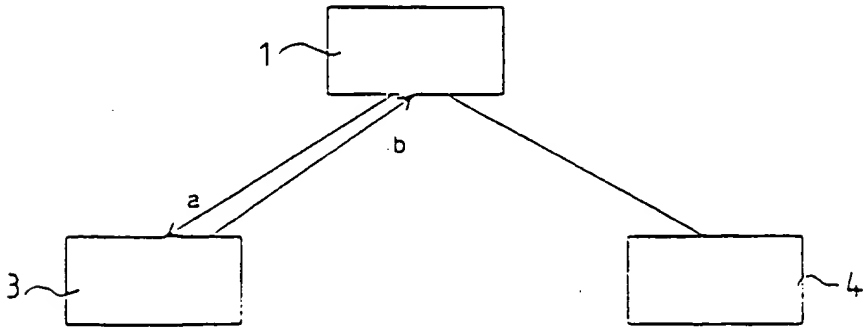


FIG.2

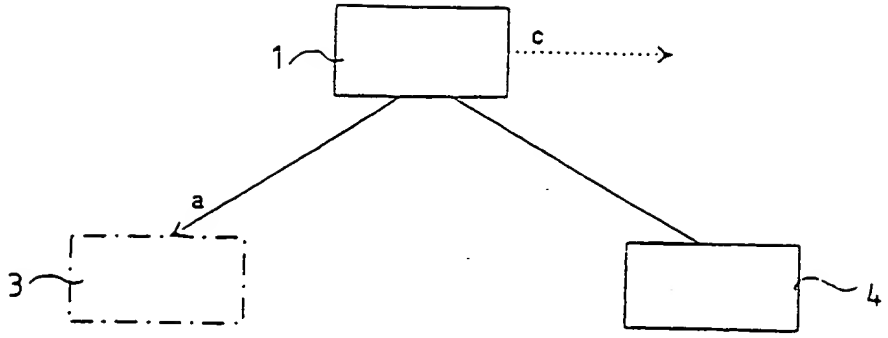


FIG. 3a

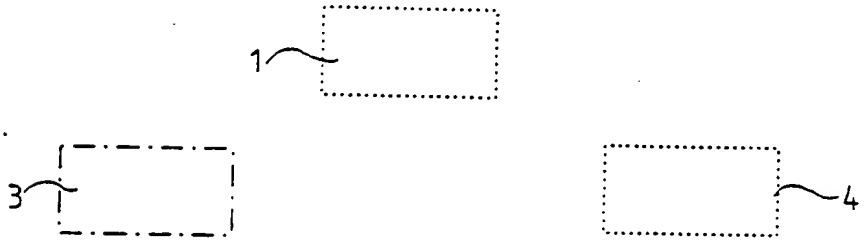


FIG. 3b

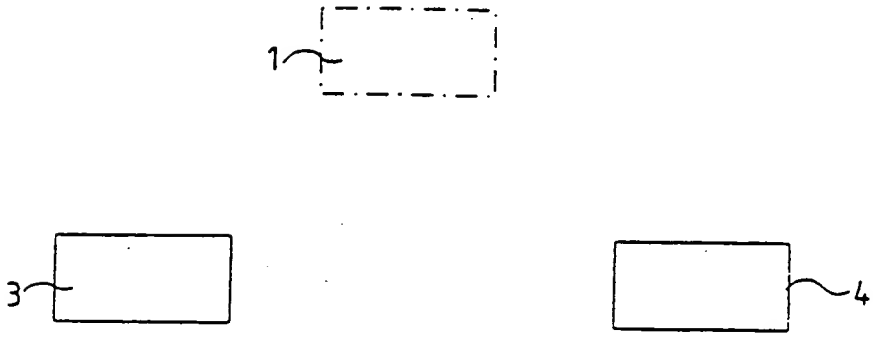


FIG. 3c

INTERNATIONAL SEARCH REPORT

International application No.
PCT/FI 93/00304

A. CLASSIFICATION OF SUBJECT MATTER

IPC5: G06F 9/46
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC5: G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

DIALOG 340, 347, 350, 351

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US, A, 4635187 (DONNA M. BARON ET AL), 6 January 1987 (06.01.87), column 2, line 5 - line 41	1-5
A	EP, A2, 0094841 (WESTERN ELECTRIC COMPANY, INCORPORATED), 23 November 1983 (23.11.83), figure 1, abstract	1-5
A	US, A, 4435766 (JUDITH G. HABER ET AL), 6 March 1984 (06.03.84), abstract	1-5

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" erlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
28 October 1993	03 -11- 1993
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86	Authorized officer Katarina Fredriksson Telephone No. +46 8 782 25 00

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT
Information on patent family members

01/10/93

International application No.
PCT/FI 93/00304

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US-A- 4635187	06/01/87	CA-A- 1218752	03/03/87
		DE-A- 3468556	11/02/88
		EP-A,B- 0165934	02/01/86
		SE-T3- 0165934	
EP-A2- 0094841	23/11/83	EP-A- 0110950	20/06/84
US-A- 4435766	06/03/84	JP-C- 1605016	13/05/91
		JP-B- 2030057	04/07/90
		JP-A- 57207957	20/12/82